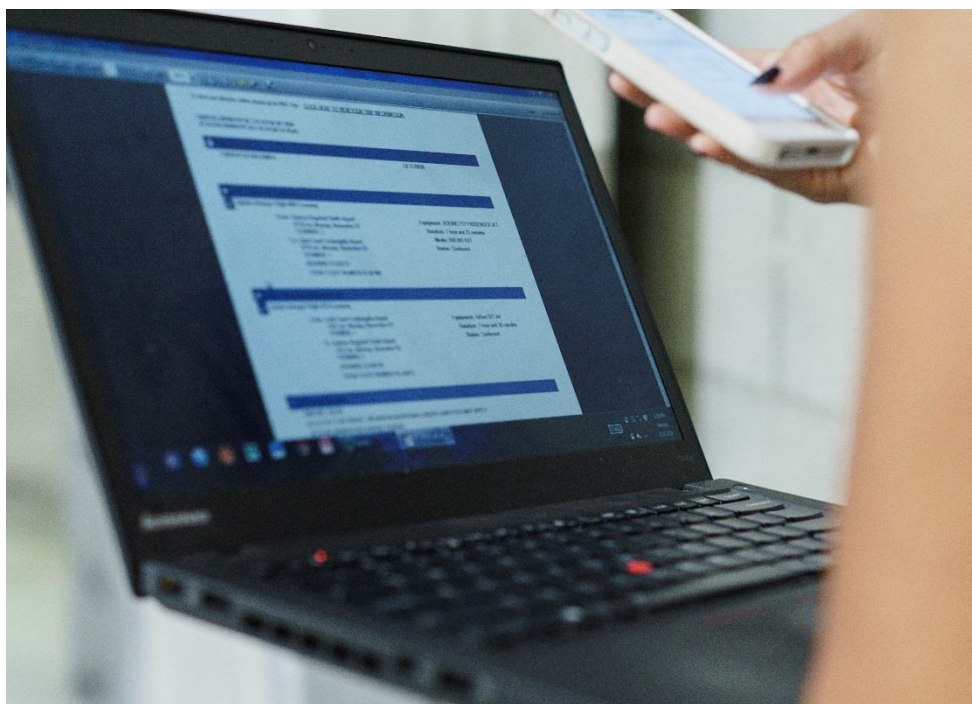


Informacinės visuomenės plėtros komitetas prie Susisiekimo ministerijos

Lietuvos atvirų duomenų formavimo metodologinių ir teisinio reglamentavimo priemonių įgyvendinimui bei tam reikalingų valstybės institucijų darbuotojų kompetencijų ugdymo paslaugos

Rekomendacijos dėl DCAT-AP naudojimo ir URI taikymo principų formuojant užklausas (R17)

2018 m. rugpjūčio 31 d.
Vilnius



Dokumento pakeitimai

Lentelė Nr. 1. Dokumento pakeitimų valdymas

Data	Aprašymas	Redaguotos dalys
2018-08-31	Pateiktas dokumento turinys derinimui	-

Informacinės visuomenės plėtros komitetui
prie Susisiekimo ministerijos
Gedimino pr. 7,
LT-01103 Vilnius

2018 m. rugpjūčio 31 d.

Projekto „Lietuvos atvirų duomenų formavimo metodologinių ir teisinio reglamentavimo priemonių įgyvendinimui bei tam reikalingų valstybės institucijų darbuotojų kompetencijų ugdymo paslaugos“ Rekomendacijos dėl DCAT-AP naudojimo ir URI taikymo principų formuojant užklausas (R17)

Gerbiamieji,

Šią Ataskaitą parengė UAB „PricewaterhouseCoopers“, (toliau – Paslaugų teikėjas arba PwC), remiantis 2017 m. gruodžio 11 d. sudaryta pirkimo sutartimi Nr. 6F-52 (toliau – Sutartis), sudaryta su Informacinės visuomenės plėtros komitetu prie Susisiekimo ministerijos (toliau – Perkančioji organizacija, IVPK).

Darbo apimtis

Rekomendacijos dėl DCAT-AP naudojimo ir URI taikymo principų formuojant užklausas (R17) (toliau – Ataskaita) apimtis buvo apibrėžta pirkimo „Lietuvos atvirų duomenų formavimo metodologinių ir teisinio reglamentavimo priemonių įgyvendinimui bei tam reikalingų valstybės institucijų darbuotojų kompetencijų ugdymo paslaugos“ techninėje specifikacijoje (toliau – Techninė specifikacija).

Darbo metodika

Ataskaita parengta remiantis Techninėje specifikacijoje nurodytais dokumentais bei reikalavimais, PwC pateiktu techniniu pasiūlymu ir susitikimų metu gauta medžiaga. PwC nesiėmė priemonių šaltinių patikimumui nustatyti ir papildomai netikrino pateiktos informacijos. Todėl PwC nesuteikia tiesioginių ar numanomų garantijų kitiems asmenims (išskyrus IVPK pagal Sutartį) dėl Ataskaitos tikslumo ar išsamumo.

Darbus pagal Sutartį atlikome iki 2018 m. rugpjūčio 31 d. Ataskaitoje neaptariami vėlesnių įvykių ar aplinkybių padariniai, taip pat vėliau paaiškėjusi informacija. Negalime nustatyti, kokią įtaką šie darbai arba tyrimai būtų turėję Ataskaitos parengimui.

Atkreipiame Jūsų dėmesį į Ataskaitoje išdėstytas svarbias mūsų pateikiamas prielaidas. Ataskaitos teikėjas neatsako kitiems asmenims (išskyrus IVPK pagal Sutartį) už Ataskaitos rengimą. Be to, Ataskaitos teikėjas neprisiima sutartinės, deliktinės ir kitokios atsakomybės (neatsižvelgiant į ieškinio formą) taikytinos teisės leistina apimtimi ir neprisiima atsakomybės už kitų asmenų patirtus padarinius (išskyrus IVPK Sutarties pagrindu) ar kitus pagal šią Programą priimtus sprendimus ar atsakymą juos priimti.

Dokumento platinimas ir atsakomybė

Asmuo, susipažinęs ir perskaitęs šią Ataskaitą, sutinka ir įsipareigoja laikytis šių sąlygų:

- Ataskaitos skaitytojas supranta, kad Ataskaita pagrįsta ekspertiniu vertinimu;
- Ataskaitos skaitytojas supranta, kad Ataskaitos teikėjo darbai yra atlikti pagal mūsų kliento – IVPK – nurodymus ir skirti tik mūsų klientui;
- Ataskaitos skaitytojas pripažįsta, kad Ataskaita parengta mūsų kliento – IVPK – nurodymu, todėl ją gali būti įtraukta ne viskas, kas svarbu skaitytojui;
- Ataskaitos skaitytojas sutinka, kad PwC, jos partneriai, įgaliotojai, darbuotojai ir atstovai neprisiima sutartinės ar deliktinės atsakomybės (įskaitant aplaidumą ir teisės aktuose nustatytų

įsipareigojimų pažeidimą ir kt.) ir neatsako už žalą, išlaidas ar nuostolius, patirtus skaitytojui vadovaujantis šiuo dokumentu ar su juo susipažinus.

Bendroji informacija

Jeigu kiltų klausimų dėl šios Ataskaitos turinio, prašome kreiptis į Justą Urboną telefonu +37065511056 arba el. paštu justas.urbonas@pwc.com.

Pagarbiai

Audrius Leipus
UAB „PricewaterhouseCoopers“
Konsultacijų skyriaus vyr. projektų vadovas

Turinys

1	Metaduomenų klasių sąvybių aprašymas.....	6
1.1	Catalogue klasė.....	6
1.2	Dataset klasė	7
1.3	Distribution klasė.....	8
1.4	Agent klasė	9
1.5	Category klasė	10
2	Metaduomenų struktūros aprašymo pavyzdžiai	11

1 Metaduomenų klasių savybių aprašymas

1.1 Catalogue klasė

Toliau lentelėje pateikiami detaliaus *Catalogue* klasės savybių aprašymus.

Lentelė Nr. 2. *Catalogue* klasės savybės

Savybė	Svarba	Aprašymas, reikšmė/pavyzdys
dataset (duomenų rinkinys)	Privaloma	Sujungia Katalogą su Duomenų rinkiniu, kuris yra Katalogo dalis. Ryšys sudaromas RDF faile, kuris yra struktūrizuotas hierarchinis failas. Pvz.: a dcat:Dataset :dataset001 ;
publisher (leidėjas)	Privaloma	Organizacija ar asmuo, kuris paskelbė Katalogą. Pvz.: dct:publisher :transparency-office ;
description (aprašymas)	Privaloma	Laisvasis tekstas, apibūdinantis Katalogą. Gali būti nurodytas keliomis kalbomis. Pvz.: dct:description ;
title (pavadinimas)	Privaloma	Katalogo pavadinimas. Gali būti pakartotinai nurodoma kitų kalbų versija. Pvz.: dct:title "Imaginary Catalog" ;
homepage (tinklapis)	Rekomenduojama	Tinklapis, kuris yra pagrindinis katalogo puslapis. Pvz.: foaf:homepage <http://example.org/catalog> ;
language (kalba)	Rekomenduojama	Kalba, kuri yra naudojama aprašant duomenų rinkinio pavadinimus, aprašymus ir kt. Gali būti kartojama, metaduomenys yra pateikiami keliomis kalbomis. Pvz.: dct:language <http://id.loc.gov/vocabulary/iso639-1/en> ;
licence (licencija)	Rekomenduojama	Licencija, pagal kuria Katalogas gali būti naudojamas ir pakartotinai naudojamas. Pvz.: dct:license rdf:resource= http://creativecommons.org/publicdomain/zero/1.0/ ;
release date (paskelbimo data)	Rekomenduojama	Oficiali katalogo išdavimo (pvz. paskelbimo) data. Pvz.: dct:issued "2011-12-05"^^xsd:date ;
themes (temos)	Rekomenduojama	Žinių organizavimo sistema, kuri yra naudojama duomenų rinkinio katalogui klasifikacijai. Pvz.: :catalog dcat:themeTaxonomy :themes . :themes a skos:ConceptScheme ; skos:prefLabel "A set of domains to classify documents" ;
update/ modification date (atnaujinimo/ pakeitimo data)	Rekomenduojama	Paskutinė data, kada Katalogas buvo keistas Pvz.: dct:modified "2011-12-05"^^xsd:date ;
has part	Neprivaloma	Nurodo susijusį katalogą, kuris yra aprašomo katalogo dalis.

Savybė	Svarba	Aprašymas, reikšmė/pavyzdys
(turi dalį)		Pvz.: dct:hasPart
is part of (yra dalis)	Neprivaloma	Nurodo susijusį katalogą, kuris yra fiziškai ar logiškai įtrauktas į aprašomą katalogą. Pvz.: dct:isPartOf
record (įrašas)	Neprivaloma	:catalog dcat:record :record-001 . :record-001 a dcat:CatalogRecord ; foaf:primaryTopic :dataset-001 ; dct:issued "2011-12-11"^^xsd:date ;
rights (teisės)	Neprivaloma	Teisių, susijusių su katalogu, nurodymas. Pvz.: dct:rights
spatial geographic (erdvinis/ geografinis)	Rekomenduojama ¹	Geografinė erdvė, kurią apima katalogo duomenys. Pvz.: dct:spatial <http://www.geonames.org/6695072> ;

1.2 Dataset klasė

Toliau esančioje lentelėje aprašomos *Dataset* klasės savybės.

Lentelė Nr. 3. *Dataset* klasės savybės

Savybė	Svarba	Aprašymas
description (aprašymas)	Privaloma	Duomenų rinkinio aprašymas tekstu.
title (pavadinimas)	Privaloma	Duomenų rinkinio pavadinimas Pvz.: dct:title "Imaginary dataset" ;
contact point (kontakciniai duomenys)	Rekomenduojama	Kontaktinė informacija, kuria galima siųsti komentarus apie duomenų rinkinį. Pvz.: dcat:contactPoint <http://example.org/transparency-office/contact> ;
dataset distribution (duomenų rinkinio platinimas)	Rekomenduojama	Susieja duomenų rinkinį su prieinamu platinimo būdu. Pvz.: dcat:distribution :dataset-001-csv ;
keyword/tag (raktinis žodis/ žyma)	Rekomenduojama	Raktiniai žodžiai ar žymos, apibūdinantys duomenų rinkinį. Pvz.: dcat:keyword "accountability","transparency" ,"payments" ;
publisher (leidėjas)	Rekomenduojama	Organizacija ar įstaiga, kuri paskelbia duomenų rinkinį. Pvz.: dct:publisher :finance-ministry ;

¹ Pakeista iš *neprivaloma* remiantis IVPK pastabomis

Savybė	Svarba	Aprašymas
theme/category (tema/ kategorija)	Rekomenduojama	Duomenų rinkinio kategorija. Duomenų rinkinys gali būti susietas su keliomis temomis. Pvz.: dcat:theme :accountability .
identifier (identifikatorius)	Neprivaloma	Pagrindinis duomenų rinkinio identifikatorius, pvz., URI ar kitas unikalus identifikatorius katalogo kontekste. Pvz.: dct:identifier "chembl"@en .
landing page (tinklapis)	Neprivaloma	Tinklapis, kuris suteikia prieigą prie duomenų rinkinio, jo platinimo ir/arba papildomos informacijos. Tinklapis turi nukreipti į pirminio duomenų teikėjo puslapį, o ne trečiosios šalies puslapį, pvz., kaupėjo. Pvz.: dcat:landingPage <http://example.org/dataset-002.html> ;
language (kalba)	Neprivaloma	Duomenų rinkinio kalba. Gali būti kartojama, jei duomenų rinkinyje naudojamos kelios kalbos. Pvz.: dct:language <http://id.loc.gov/vocabulary/iso639-1/en> ;
spatial/ geographical coverage (erdvinis/ geografinis paplitimas)	Neprivaloma	Geografinis regionas, kurį apima duomenų rinkinys. Pvz.: dct:spatial <http://www.geonames.org/6695072> ;
temporal coverage (laikotarpis)	Neprivaloma	Laikotarpis, kurį apima duomenų rinkinys. Pvz.: dct:temporal <http://reference.data.gov.uk/id/quarter/2006-Q1> ;
type (tipas)	Neprivaloma	Duomenų rinkinio tipas. Kontroliuojamo žodyno reikšmės dar nėra nustatytos.
update/ modification date (atnaujinimo/ pakeitimo data)	Rekomenduojama ²	Paskutinė data, kada duomenų rinkinys buvo keistas. Pvz.: dct:modified "2011-12-05"^^xsd:date ;

1.3 Distribution klasė

Toliau esančioje lentelėje aprašomos *Distribution* klasės savybės.

Lentelė Nr. 4. Distribution klasės savybės

Savybė	Svarba	Aprašymas
access URL (prieigos URL)	Privaloma	URL nuoroda, kuri suteikia prieigą prie duomenų rinkinio platinimo. Gali būti nurodoma tokia informacija, kaip gauti duomenį rinkinį. Pvz.: dcat:accessURL <http://example.org/dataset-002.html> ;

² Pakeista iš *neprivaloma*

Savybė	Svarba	Aprašymas
description (aprašymas)	Rekomenduojama	Platinimo aprašymas tekstu. Gali būti kartojamas keliomis kalbomis.
format (formatas)	Rekomenduojama	Platinimo failo formatas. Pvz.: dct:format "text/html";
licence (licencija)	Rekomenduojama	Licencija, pagal kurią prieinamas platinimas. Pvz.: dct:license rdf:resource= http://creativecommons.org/publicdomain/zero/1.0/ ;
byte size (dydis baitais)	Neprivaloma	Platinimo dydis, nurodytas baitais. Pvz.: dcat:byteSize "5120"^^xsd:decimal ;
download URL (parsisiuntimo nuoroda)	Neprivaloma	URL nuorodą į parsisiunčiamą failą nurodytu formatu. Pvz.: dcat:downloadURL < http://www.example.org/files/001.csv > ;
media type (medijos tipas)	Rekomenduojama ³	Platinimo medijos tipas, kaip nurodyta oficialiame IANA medijų tipų registre. Pvz.: dcat:mediaType "text/csv" ;
release date (paskelbimo data)	Neprivaloma	Oficiali platinimo išdavimo (pvz. paskelbimo) data. Pvz.: dct:issued "2011-12-11"^^xsd:date ;
rights (teisės)	Neprivaloma	Teisių, susijusių su platinimu, pareiškimas. Pvz.: dct:rights
title (pavadinimas)	Neprivaloma	Platinimo pavadinimas. Gali būti kartojamas keliomis kalbomis. Pvz.: dct:title "ChEMBL" ;
update/ modification date (atnaujinimo/ pakeitimo data)	Neprivaloma	Paskutinė data, kada platinimas buvo keistas. Pvz.: dct:modified "2011-12-05"^^xsd:date ;

1.4 Agent klasė

Toliau esančioje lentelėje aprašomos *Agent* klasės savybės.

Lentelė Nr. 5. *Agent* klasės savybės

Savybė	Svarba	Aprašymas
name (vardas)	Privaloma	Agento pavadinimas. Gali būti kartojamas keliomis kalbomis. Pvz.: <foaf:name>Peter Parker</foaf:name>
type (tipas)	Rekomenduojama	Agento tipas, kuris leidžia katalogą ar duomenų rinkinį. Pvz.: dct:type

³ Pakeista iš *neprivaloma*

1.5 Category klasė⁴

Toliau esančioje lentelėje aprašomos *Category* klasės savybė.

Lentelė Nr. 6. *Category* klasės savybė

Savybė	Svarba	Aprašymas
preferred label (pageidaujama žyma)	Privaloma	Pageidaujama kategorijos etiketė. Gali būti kartojama keliomis kalbomis. Pvz.: <skos:prefLabel xml:lang="en">Access to countryside</skos:prefLabel>

⁴ Klasė „Kategorija“ apibrėžta Europos Komisijos standarte, kaip aprašyta čia. Vis dėlto W3C specifikacijoje pateikiamas kitas kategorijų valdymas, naudojant klases "concept skos:Concept" ir "skos:ConceptScheme", pakeičiant šias klases. Kadangi oficiali šių klasių integracija nėra prieinama Joinup platformoje, rekomenduoja laikytis standarto, iki tolesnio pranešimo ar atnaujinimo.

2 Metaduomenų struktūros aprašymo pavyzdžiai

Failo nuskaitymui, RDF faile turėtų būti nurodyti šie elementai:

```
xmlns:time="http://www.w3.org/2006/time#"
xmlns:dct="http://purl.org/dc/terms/"
xmlns:dcat="http://www.w3.org/ns/dcat#"
xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
xmlns:tema="http://datos.gob.es/kos/sector-publico/sector/"
xmlns:auto="http://datos.gob.es/recurso/sector-publico/territorio/Autonomia/"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
https://github.com/ckan/ckanext-dcat/blob/master/examples/catalog.rdf
```

RDF failo pateikimo pavyzdys:

```
<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF
  xmlns:time="http://www.w3.org/2006/time#"
  xmlns:dct="http://purl.org/dc/terms/"
  xmlns:dcat="http://www.w3.org/ns/dcat#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:tema="http://datos.gob.es/kos/sector-publico/sector/"
  xmlns:auto="http://datos.gob.es/recurso/sector-publico/territorio/Autonomia/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">

  <dcat:Catalog rdf:about="https://data.some.org/catalog">
    <dct:identifier>https://data.some.org/catalog</dct:identifier>
    <dct:title>An example catalog</dct:title>
    <dct:description>A longer description of what the catalog is, who maintains it, etc. Special characters
    should be properly handed (if the file is utf8 encoded), for instance for catalogs from Sant
    Adrià.</dct:description>
    <dct:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2013-04-
    26T00:00:00Z</dct:issued>
    <dct:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2013-05-
    22T09:38:45Z</dct:modified>
    <dct:language>en</dct:language>
    <foaf:homepage rdf:resource="https://data.some.org/the/actual/catalog"/>
    <dct:license rdf:resource="https://data.some.org/link/to/license"/>
    <dcat:dataset>
      <dcat:Dataset rdf:about="https://data.some.org/catalog/datasets/1">
        <dct:identifier>https://data.some.org/catalog/datasets/1</dct:identifier>
        <dct:title>Example dataset 1</dct:title>
        <dct:description>A longer description of dataset 1: includes three distributions, a CSV file, a PDF file
        and an API. Some special characters: Θεσσαλονίκη.</dct:description>
        <dcat:keyword>stats</dcat:keyword>
        <dcat:keyword>pollution</dcat:keyword>
        <dct:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2013-04-
        26T00:00:00Z</dct:issued>
        <dct:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2013-04-
        26T00:00:00Z</dct:modified>
        <dct:language>en</dct:language>
```

```

< dct:license rdf:resource="https://data.some.org/link/to/license"/>
< dct:publisher>
  < foaf:Organization>
    < foaf:name>Publishing Organization for dataset 1</foaf:name>
    < foaf:mbox>contact@some.org</foaf:mbox>
  </foaf:Organization>
</dct:publisher>
< dcat:distribution>
  < dcat:Distribution>
    < dct:identifier>https://data.some.org/catalog/datasets/1/distribution/1</dct:identifier>
    < dct:title>A distribution for this dataset in CSV format</dct:title>
    < dct:description>A longer description for the distribution</dct:description>
    < dcat:downloadURL
rdf:datatype="http://www.w3.org/2001/XMLSchema#anyURI">https://data.some.org/the/actual/catalog/datasets/1/resources/1.csv</dcat:downloadURL>
    < dcat:mediaType>text/csv</dcat:mediaType>
    < dct:license rdf:resource="https://data.some.org/link/to/license"/>
  </dcat:Distribution>
</dcat:distribution>
< dcat:distribution>
  < dcat:Distribution>
    < dct:identifier>https://data.some.org/catalog/datasets/1/distribution/2</dct:identifier>
    < dct:title>A distribution for this dataset in PDF format</dct:title>
    < dct:description>A longer description for the distribution</dct:description>
    < dcat:downloadURL
rdf:datatype="http://www.w3.org/2001/XMLSchema#anyURI">https://data.some.org/the/actual/catalog/datasets/1/resources/2.pdf</dcat:downloadURL>
    < dct:format>
      < dct:IMT>
        < rdf:value>application/pdf</rdf:value>
      </dct:IMT>
    </dct:format>
    < dct:license rdf:resource="https://data.some.org/link/to/license"/>
  </dcat:Distribution>
</dcat:distribution>
< dcat:distribution>
  < dcat:Distribution>
    < dct:identifier>https://data.some.org/catalog/datasets/1/distribution/3</dct:identifier>
    < dct:title>A distribution for this dataset as an API</dct:title>
    < dct:description>A longer description for the distribution</dct:description>
    < dcat:accessURL
rdf:datatype="http://www.w3.org/2001/XMLSchema#anyURI">https://data.some.org/the/actual/catalog/link/to/an/api</dcat:accessURL>
  </dcat:Distribution>
</dcat:distribution>
</dcat:Dataset>
</dcat:dataset>
</dcat:Catalog>
</rdf:RDF>

```

Toliau nurodomas metaduomenų turinys pateiktas JSON formatu⁵:

⁵ Šaltinis: http://guidance.data.gov.uk/dcat_fields.html

```
[
  {"title": "Live traffic information from the Highways Agency",
    "license": "No license provided",
    "publisher": {"name": "Highways Agency", "mbox": "test@test.com"},
    "distribution": [
      {
        "downloadURL": "https://s3-eu-west-
1.amazonaws.com/lmtesting2810/HATRIS_15MinuteMIDAS_YYYY-MM-DD_o.csv.zip",
        "title": "Hattris 15 Min Midas o",
        "format": "application/zip"
      }
    ]
  },
  {"title": "Roadworks locations",
    ...
  }
]
```

Pastaba: W3C pateikia daugiau informacijos apie JSON naudojimą čia: <https://www.w3.org/TR/json-ld/>.